CrossMark

Stochastic optimization models in forest planning: a progressive hedging solution approach

**Fernando Badilla Veliz · Jean-Paul Watson · Andres Weintraub ·
Roger J.-B. Wets · David L. Woodruff**

Abstract We consider the important problem of medium term forest planning with an integrated approach considering both harvesting and road construction decisions in the presence of uncertainty modeled as a multi-stage problem. We give strengthening methods that enable the solution of problems with many more scenarios than previously reported in the literature. Furthermore, we demonstrate that a scenario-based decomposition method (Progressive Hedging) is competitive with direct solution of the extensive form, even on a serial computer. Computational results based on a real-world example are presented.

1 Introduction

In this paper, we consider the problem of medium term forest planning. A deterministic version of this problem has been proposed and solved for a real Chilean forest firm (Andalaft et al. 2003). The main decisions to be made involve which units (lots) to harvest in each period of the planning horizon and what roads need to be built in order to access the harvest units. This is a well-known problem, with difficulties for large-scale problems, due to computationally inefficient MIP formulations. Although current versions of CPLEX can solve the problem

F. B. Veliz · A. Weintraub
Departmento de Ingenieria Industrial, Universidad de Chile, Santiago, Chile

J.-P. Watson
Discrete Math and Complex Systems Department, Albuquerque, NM, USA

R. J.-B. Wets
Department of Mathematics, University of California Davis, Davis, CA, USA

D. L. Woodruff (✉)
Graduate School of Management, University of California Davis, Davis, CA, USA
e-mail: dlwoodruff@ucdavis.edu

directly, Andalaft et al. (2003) introduce a solution approach based on Lagrangian relaxation and a strengthening of the LP formulation, enabling solution of an instance with 17 forests, which are linked by demand constraints at the firm level. This problem was solved considering deterministic demand and price conditions for each period. Quinteros et al. (2009) consider a variant of this problem with market uncertainty, represented through 16 sampled scenarios with both timber prices and upper and lower bounds on demand in each period. In this work each scenario represents all uncertain parameters in all periods, from period 1 to the final time period in the planning horizon. Periods in this problem were defined as semesters, to account for seasonal effects. Given the increased difficulty of solving stochastic models relative to their deterministic counterparts, the paper considered only one forest, the largest of the original 17.

The stochastic model developed by Quinteros et al. (2009) considered the objective of maximizing expected net present value, subject to the solution being feasible under all scenarios. This is a standard treatment of uncertainty in the objective of a stochastic optimization model. The constraints in this case need to include the well-known non-anticipativity constraints (Wets 1975), which state that, given 2 scenarios S_i and S_k, if the representation (i.e., data) of scenarios S_i and S_k are identical up to period t, then the decisions taken in scenarios S_i and S_k must also be identical up to period t.

The non-anticipativity constraints in a stochastic optimization model can be expressed in two ways, to create a so-called extensive form. In an explicit formulation, the non-anticipativity constraints relate scenario-specific instances of the decision variables at each node in the scenario tree. In an implicit formulation, the non-anticipativity constraints are expressed by introducing shared, common decision variables at each node in the scenario tree. The implicit formulation is more compact and thus can lead to problems that are easier to solve (Quinteros et al. 2009). However, it is often difficult to solve stochastic optimization instances, even with an implicit formulation, given their typically large dimension. In Quinteros et al. (2009), the stochastic forest planning problem was solved by incorporating the non-anticipativity constraints in an implicit manner, by defining one path through the scenario tree for each scenario, but with variables shared by the all scenarios at a node of the tree. By branching in a coordinated way along all paths, the non-anticipativity constraints are satisfied automatically. Alonso-Ayuso et al. (2003) first developed this solution approach. This type of decomposition enabled more efficient solution of the stochastic optimization model. In Quinteros et al. (2009), a problem with 16 scenarios was solved considering several sources of uncertainty. In this paper, we present a different approach to handle the non-anticipativity constraints, to enable computationally efficient consideration of even larger numbers of scenarios.

We use the algorithm Progressive Hedging (PH) (Rockafellar and Wets 1991), which is based on the separation of the problem by scenarios, thus eliminating the need to explicitly introduce non-anticipativity constraints. In PH, each scenario is solved independently and iteratively. Naturally, the solutions obtained under each scenario do not initially satisfy non-anticipativity, and variables that should be equal by non-anticipativity take different values. The PH algorithm attempts to equalize the values of these variables by penalizing deviations from the average value in all scenarios, through augmentation of the objective function. In an iterative way, convergence to a solution ultimately satisfying the non-anticipativity constraints (without their representation, either in explicit or implicit form) can be developed. The advantage of PH lies in the possibility of solving problems with very large numbers of scenarios, such that an extensive form of the stochastic optimization problem is either too difficult to solve directly, requires too much memory (e.g., in the course of a branch-and-cut procedure), or both. PH and a few other methods have been proposed for multi-stage

stochastic programming with integers in all stages, but applications are relatively scarce. For alternatives to PH, see, e.g., Caroe and Schultz (1999), Römisch and Schultz (2001), Guglielmo and Sen (2004), Escudero et al. (2009).

In this paper, we solve the same basic stochastic forest planning problem as (Quinteros et al. 2009). However, in order to render the model more realistic—in particular, to allow for a wide variety of potential undesirable scenarios—we focus on the need to consider a substantially larger number of scenarios. The scenarios for our test case additionally include uncertainty in timber growth and yield.

Forest planning problems have been solved over the past few decades using integer programming models, including an integrated approach that considers both harvesting and road construction decisions (Andalaft et al. 2003; Weintraub et al. 1994, 1995; Kirby et al. 1986, Weintraub and Navon 1976). Introducing binary decision variables into forest planning problems has become integral in the course of making these models more realistic over time. There are two components of these models where discrete decision have been widely leveraged: (a) Road building and (b) All-or-nothing decisions on harvesting areas (lots). Such discrete decisions warrant careful consideration, as they induce mixed-integer optimization models, which can be very difficult to solve in practice. Because our model considers such features, we now briefly review the literature related to the use of discrete decision variables in forest planning.

Road building decisions: Between the late 1970's and 1990's, the US Forest Service introduced explicit decision making related to road building (Weintraub and Navon 1976; Kirby et al. 1986). Originally, harvesting models were modeled as continuous optimization problems, which were solved independently of road building models. Jones et al. (1986) provide a thorough analysis showing the significant advantages of integrating the two types of decisions. Weintraub et al. (1995) describe an application involving several regions of the US Forest Service, in which planning was carried out using a mixed-integer programming (MIP) model to integrate road building and harvesting decisions. The primary challenge in this problem involved the fixed cost network flow portion of the model.

Given the computational power and algorithmic advances available at the time, heuristics were needed to find approximate solutions. Guignard et al. (1994) discussed model tightening to better solve these MIP problems. In Andalaft et al. (2003), a similar problem was applied to Forestal Millalemu, a forest company in Chile, which was used again for integrated planning of harvesting and road building decisions. In this case, the company managed 17 separate forests, linked by demand constraints. CPLEX 3.0 was used to solve the problem, and obtained a solution with a optimality bound of about 20 % in 6 h. Using Lagrangian relaxation and a strengthening of the formulation, a solution within 1 % of optimal was obtained in one hour. However, the same problem run years later, with CPLEX 9.0, could be solved to optimality in minutes. This paper presents a stochastic version of the (Andalaft et al. 2003) model, applied to one of the 17 forests (the largest) of the Forestal Millalemu problem. MIP models for forest road building problems have been solved by researchers in Scandinavia, in this case for road upgrade decisions (Frisk et al. 2006).

Harvesting decisions: Harvesting decisions are naturally modeled using 0–1 variables. The constraint that a cutting unit be harvested completely or not at all has been particularly important for environmental reasons, where adjacent units cannot be harvested in the same period. This constraint is based on what is referred to as the maximum opening size, which states that there is a maximum contiguous area (typically about 40 has). Contributing factors include wildlife concerns (elks will not feed in a clearing unless they are relatively near to the protection provided by grown trees), scenic beauty, and erosion factors. Thus, in the planning process, units are defined to be no larger than this maxi-

mum opening size area. Further, if one unit k is harvested, the neighboring units cannot be harvested until trees in unit k have reached a certain height. While in this paper we do not impose these adjacency constraints, it is important to have this option, which is now being applied in a deterministic problem. Second, for production and logistical reasons, given the size of the units, it makes economic sense that once a unit is intervened, it should be completely harvested. These concepts have proved critical, and have been used in forest planning for over 20 years. Additional conditions are often imposed, such as needing also to maintain large contiguous areas of mature trees for wildlife protection. In many cases, heuristics have been used to solve these problems (Murray and Church 1995). Exact algorithms have also been proposed for this problem class. However, for larger adjacency problems, even the latest version of CPLEX cannot solve these problems. More complex algorithms, including column generation, have performed better (Goycoolea et al. 2005; Richards and Gunn 2000; Constantino et al. 2008). Goycoolea et al. (2009) present a comparison of these methods. Another approach is presented by McNaughton and Ryan (2008).

The remainder of this paper is organized as follows. Section 2 presents the deterministic medium term forest planning model to be solved. Section 3 describes the uncertainties involved in our stochastic optimization model. Section 4 details the stochastic optimization model. In Sect. 5, we describe the PH algorithm and our specific implementation. Section 6 introduces and discusses the computational results. We conclude in Sect. 7 with a summary of our contributions.

2 The forest planning optimization model

We consider a tactical (medium-term) planning model first developed for a Chilean forest firm, Millalemu. The original problem considered the planning of 17 forests, spatially separated. The planning process horizon was 3 years, divided into summer and winter periods, to account for seasonal effects. The planning decisions involved which units to harvest in each period, and what roads to build to access the harvesting units not already connected to main public roads.

Additionally, stocking areas were used to store timber from summer to winter, and timber flows defined the movement of logs from origins in the forest to destinations. This problem was introduced by Andalaft et al. (2003), and was solved by using Lagrangian relaxation to decouple the forests—which are linked by global demand constraints. The addition of strengthening and lifting constraints to the initial formulation yielded significant improvements in the computational efficiency of the mixed-integer optimization solves. Such enhancements resulted in reductions of the optimality gap for the most difficult data set tested from 162 to 1.6 %, and for all data sets to 2.6 % or less.

In this paper, we work with a simplified version of the original deterministic optimization model. The additional complexities would only add to the computational burden, and are not needed to analyze the complexities of the stochastic optimization model and solution procedure. The specific simplifications we consider are as follows:

(a) We only consider one type of timber, eliminating the possibility of downgrading (cutting) timber from higher diameters to lower diameters.
(b) We do not consider stocking yards, eliminating the storing of timber from summer periods to winter periods.
(c) We only define one type of road, thus eliminating the upgrading of roads.
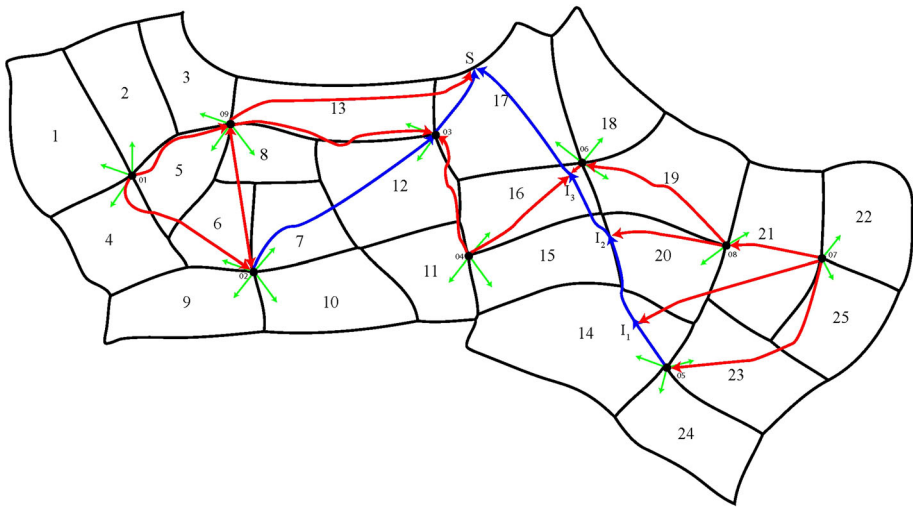
Fig. 1 Diagram of the Los Copihues forest

Figure 1 shows a description of the forest we will analyze, consisting of 25 units to be harvested. Existing roads are depicted as blue arrows, potential roads are depicted as red arrows, and the access to each unit is depicted by green arrows. Finally, the point S represents the exit node of the forest, which links with a public road.

For the sake of brevity, we do not present the simplified deterministic optimization model here. Rather, we introduce the full stochastic optimization model in Sect. 4, with a scenario-based formulation, driven by parameter uncertainty as described next in Sect. 3.

3 Uncertainties in forest planning

Typical uncertainties occurring in forestry planning are documented by Martell et al. (1998), and summarized as follows:

(a) **Market uncertainties**: These uncertainties are expressed mainly through prices for the different timber products, and in some cases also through bounds on demand. Lower demand bounds can reflect minimum amounts a firm needs to sell to keep a market, while upper demand bounds can reflect the maximum amount possible to sell in a market. These properties are a natural result of varying global market conditions as well as local realities in countries where the firm exports. Chile exports almost 90 % of its lumber and pulp products, so its forest firms are heavily dependent on global economic conditions.

(b) **Natural variations in future growth and yields**: These uncertainties address the cubic volume (m^3) per hectare that will be harvested in the future from each unit. Regression and simulation models allow reasonable point predictions of the values of future volumes that will be harvested, but there is a variability that can be modeled though approximate probability functions, with estimates of the averages and standard deviations. One of the approaches that has been proposed is to incorporate this type of variability is chance constraint programming (Hof and Pickens 1991; Weintraub and Vera 1991), in the context of long range planning LP models. Note that if integer variables are incorporated,

as is the case in our model, which includes road building, this approach is no longer a computationally practical alternative.

(c) **The effect of fires or pests**. These uncertainties model the impact of fires and/or pests on the harvesting of particular units in each period, as described in Martell (2007).

We consider the first two types of uncertainties, expressed as scenarios that specify value for the uncertain parameters in each period through the planning horizon. This is a well known form for expressing uncertainty, in the form of a (sampled) scenario tree.

As discussed in Sect. 2, for the purpose of developing the stochastic model we will consider only one forest, the largest one of the 17 in the original problem, as was done by Quinteros et al. (2009). Details of the scenario generation process are provided in Sect. 6.

4 The stochastic optimization model

We present now our stochastic optimization model, where the deterministic optimization model is adapted for each scenario, and the scenarios are coupled by non-anticipativity constraints.

Sets

T : *Time periods in the planning horizon*, $\{t\}$
O : *Origin nodes of harvests*, $\{o\}$.
J : *Road intersection nodes*, $\{j\}$
E : *Timber exit nodes*, $\{e\}$
H : *Harvest units*, $\{h\}$
H_o : *Units adjacent to origin node* $\{o\}$
O_h : *Origin nodes adjacent to unit* $\{h\}$
R^E : *Existing roads, arcs* $\{k, l\}$
R^P : *Potential roads, arcs* $\{k, l\}$
$R = R^E \cup R^P$: *All roads. arcs* $\{k, l\}$
S : *Scenarios* $\{s\}$

Deterministic parameters

A_h : *Area of unit h* (ha)
$U_{k,l}^t$: *flow capacity on road* (k, l) *in period t* (m^3)
P_h^t : *Cost of harvesting one ha of unit h in period t* (USD/Ha)
Q_o^t : *Cost of producing one* m^3 *at origin o in period t* (USD/m^3)
$C_{k,l}^t$: *Cost of building road* (k, l) *in period t* (USD)
$D_{k,l}^t$: *Cost of hauling one* m^3 *on road* (k, l) *in period t* (USD/m^3)
$\Pr(s)$: *Probability of the occurence of scenario s*
d^t : *Discount factor for period t*

Scenario-dependent parameters

$R_e^{t,s}$: *Sale price in period t under scenario s* (USD/m^3)
$\underline{Z}^{t,s}, \bar{Z}^{t,s}$: *Lower and upper bounds on demand under scenario s in period t* (m^3)
$a_h^{t,s}$: *Productivity of unit h in period t under scenario s* (m^3/ha)

Decision variables

$\delta_h^{t,s} \in \{0, 1\} : \begin{cases} 1 \ if \ unit \ h \ is \ harvested \ in \ period \ t \ under \ scenario \ s \\ 0 \ otherwise \end{cases}$

$\gamma_{k,l}^{t,s} \in \{0, 1\} : \begin{cases} 1 \ if \ road \ (k, l) \ is \ built \ in \ period \ t \ under \ scenario \ s \\ 0 \ otherwise \end{cases}$

$f_{k,l}^{t,s} \geq 0 : Flow \ of \ timber \ transported \ on \ arc \ (k, l) \ in \ period \ t \ under \ scenario \ s \ (\text{m}^3)$

$z_e^{t,s} \geq 0 : Timber \ sold \ in \ period \ t \ at \ exit \ e \ under \ scenario \ s \ (\text{m}^3)$

Constraints

1. *Balance of flow at network nodes*
(*a*) *At harvest origin nodes*

$$\sum_{h \in H_0} a_h^{t,s} A_h \delta_h^{t,s} + \sum_{(k,o) \in K} f_{k,0}^{t,s} - \sum_{(o,k) \in K} f_{0,k}^{t,s} = 0 \ \forall o \in O, t \in T, s \in$$

(*b*) *At road intersection nodes*

$$\sum_{(k,j) \in K} f_{k,j}^{t,s} - \sum_{(j,k) \in K} f_{j,k}^{t,s} = 0 \quad \forall t \in T, s \in S, j \in J$$

(*c*) *At exit nodes*

$$z_e^{t,s} = \sum_{(k,e) \in K} f_{k,e}^{t,s} - \sum_{(e,k) \in K} f_{e,k}^{t,s} \quad \forall e \in E, t \in T, s \in S$$

(*d*) *No stocks are held between periods*

$$\sum_{h \in H} a_h^{t,s} A_h \delta_h^{t,s} - \sum_{e \in E} z_e^{t,s} = 0 \quad \forall t \in T, s \in S$$

2. *Bounds for timber production*

$$\underline{Z}^{t,s} \leq \sum_{e \in E} z_e^{t,s} \leq \bar{Z}^{t,s}$$

3. *Building of potencial roads*
(*a*) *Flow up to capacity if road* (*k, l*) *is built*

$$f_{k,l}^{t,s} \leq U_{k,l}^{t,s} \sum_{1 \leq m \leq t} \gamma_{k,l}^{m,s} \quad \forall (k, l) \in R^P, t \in T, s \in S$$

(*b*) *Roads are built at most once*

$$\sum_{t \in T} \gamma_{k,l}^{t,s} \leq 1 \quad \forall (k, l) \in R^P, s \in S$$

4. *Capacity of existing roads*
(*a*) *Respect the capacity of each road*

$$f_{k,l}^{t,s} \leq U_{k,l}^{t,s} \quad \forall (k, l) \in R^E, t \in T, s \in S$$

(*b*) *Upper bound of flow capacity for all roads*

$$f_{k,l}^{t,s} \leq \sum_{w \in O} \sum_{h \in H_w} a_h^{t,s} A_h \delta_h^{t,s} \quad \forall (k, l) \in K, t \in T, s \in S$$

5. *Each cell is harvested at most once*

$$\sum_{t \in T} \delta_h^{t,s} \leq 1 \qquad \forall h \in H,\, s \in S$$

Objective function

$$\max \sum_{s \in S} \Pr(s)\left(TS^s - HC^s - PC^s - RC^s - TC^s\right)$$

where:

$$TS^s = \sum_{e \in E}\sum_{t \in T} d^t R_e^{t,s} z_e^{t,s} \qquad \forall s \in S$$

$$HC^s = \sum_{h \in H}\sum_{t \in T} d^t P_h^{t,s} A_h \delta_h^{t,s} \qquad \forall s \in S$$

$$PC^s = \sum_{o \in O}\sum_{t \in T} d^t Q_o^t \left(\sum_{h \in H_o} a_h^{t,s} A_h \delta_h^{t,s}\right) \qquad \forall s \in S$$

$$RC^s = \sum_{(k,l) \in R^P}\sum_{t \in T} d^t C_{k,l}^{t,s} \gamma_{k,l}^{t,s} \qquad \forall s \in S$$

$$TC^s = \sum_{(k,l) \in K}\sum_{t \in T} d^t D_{k,l}^{t,s} f_{k,l}^{t,s} \qquad \forall s \in S$$

The objective function specifies maximization of the net revenue, which is composed of the income from timber sales (TS) and four types of expenses: the harvesting cost (HC), the cost of the production at the origin nodes (PC), the road construction cost (RC), and the transportation costs (TC).

Additional constraints

1. *Direct strengthenings*
   We additionally include several strengthening constraints to improve mixed-integer solve times, first introduced by Andalaft et al. (2003). The following notation is used to express these constraints:

$$SC = \{(i,j) \in R^P\,/\,[(i,\cdot)\bigvee(\cdot,i)\in R^E \wedge j \in O]\bigvee[(j,\cdot)\bigvee(\cdot,j)\in R^E \bigwedge i \in O]\}$$
$$N(k,l) = \{(i,j) \in R^P\,/\,i = k,l \bigvee j = l,k\}$$

   The specific constraints employed are given as follows:
   (a) *Do not build isolated roads*

$$\gamma_{k,l}^{t,s} \leq \sum_{q \leq t}\sum_{(i,j) \in N(k,l)} \gamma_{i,j}^{q,s} \qquad \forall t \in T,\, (k,l) \in R^P \backslash SC,\, s \in S$$

   (b) *Do not harvest units unless a connecting road has been built*

$$\delta_h^{t,s} \leq \sum_{q \leq t}\sum_{(i,j) \in N(h)} \gamma_{i,j}^{q,s} \qquad \forall t \in T,\, s \in S,\, h\ not\ connected\ to\ an\ existing\ road$$

   where $N(h) = \left\{(i,j) \in R^P\,|\,i \in O_h \bigvee j \in O_h\right\}$. We observe that *h not connected to an existing road* $\Leftrightarrow \{(i,j) \in R^E\,|\,i \in O_h \bigvee j \in O_h\} = \emptyset$.
2. *Liftings*

(a) *Lifting on road capacities*, where flows in the LHS on all periods up to t are added up, considering that timber can be harvested at most once

$$\sum_{g \leq t} f_{k,l}^{g,s} \leq U_{k,l}^{t,s} \sum_{1 \leq m \leq t} \gamma_{k,l}^{m,s} \quad \forall \, (k,l) \, \epsilon \, R^P, \, t \, \epsilon \, T, \, s \, \epsilon \, S$$

$$\sum_{g \leq t} f_{k,l}^{g,s} \leq U_{k,l}^{t,s} \quad \forall \, (k,l) \, \epsilon \, R^E, \, t \, \epsilon \, T, \, s \, \epsilon \, S$$

(b) *Lifting on not building isolated roads*

$$\sum_{g \leq t} \gamma_{k,l}^{g,s} \leq \sum_{g \leq t} \sum_{(i,j) \, \epsilon \, N(k,l)} \gamma_{i,j}^{g,s} \quad \forall \, (k,l) \, \epsilon \, R^P \backslash SC, \, t \, \epsilon \, T, \, s \, \epsilon \, S$$

(c) *Lifting on the constraints of not harvesting unconnected units*

$$\sum_{g \leq t} \delta_h^{g,s} \leq \sum_{g \leq t} \sum_{(i,j) \, \epsilon \, N(h)} \gamma_{i,j}^{g,s}$$

$$\forall \, t \, \epsilon \, T, \, s \, \epsilon \, S, \, h \, not \, connected \, to \, an \, existing \, road$$

3. *New Strengthenings*

(a) *Flow balance in the network*

$$\sum_{h \in H} a_h^{t,s} A_h \delta_h^{t,s} - \sum_{e \in E} z_e^{t,s} = 0 \quad \forall \, t \, \epsilon \, T, \, s \, \epsilon \, S$$

(b) *Road flow capacity*

$$f_{k,l}^t \leq \sum_{h \in H} a_h^{t,s} A_h \delta_h^{t,s} \quad \forall \, (k,l) \in R, \, t \, \epsilon \, T, \, s \, \epsilon \, S$$

The new strengthening provided significant improvement in the efficiency of the runs.

5 The progressive hedging algorithm

Consider a stochastic optimization problem such as that presented in Sect. 4, where scenarios have been pre-defined. Figure 2 presents an example problem with 4 time periods and 5 scenarios. Figure 2a shows the scenario tree in a standard form, while Fig. 2b shows the 5 scenarios in a disaggregated form, in which the non-anticipativity constraints on decision variables are represented by the circling of nodes. In period 1, all scenarios share identical decisions, as no knowledge of future uncertainty has been acquired yet. In contrast, we see in period 2 that node 2 represents scenarios 1, 2 and 3, while node 3 shares scenarios 4 and 5.

As the uncertainties can take only a finite number of values when a scenario tree is provided as input, the stochastic optimization problem can be stated as follows:

$$\min_x \sum_{s \in S} \Pr(s) f(x, s)$$
$$s.t. x_s \in C_s \quad \forall s \in S$$
$$x \in \mathcal{N}$$

The model seeks to minimize the expected cost across all scenarios, weighted by their probabilities Pr(s). Solutions must satisfy feasibility for all scenarios when taken independently,
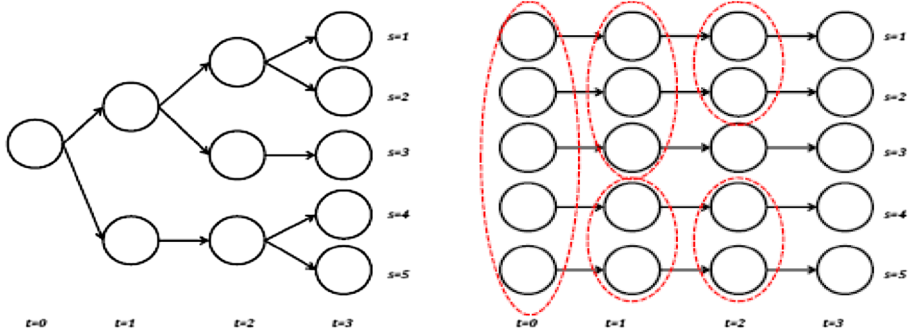
Fig. 2 Example of a scenario tree in standard (*left* figure) and disaggregated (*right* figure) forms

and must satisfy the non-anticipativity constraints at each node in the scenario tree when taken together. We use C_s to represent the constraints for each scenario. We use $\mathcal{N}$ to represent the set of non-anticipative solution vectors, i.e., those for which scenarios that share a node in the scenario tree have the same value for all those decision vector elements associated with that node..

The basic idea of the PH algorithm is to relax the non-anticipativity constraints and solve the scenario sub-problems independently. Solving these problems separately is naturally much easier than solving the extensive form directly. However, the non-anticipativity constraints will be satisfied only in very rare cases. That is, the values of the variables which should be equal to satisfy non-anticipativity are not equal. The PH algorithm then iteratively solves the scenario sub-problems, gradually enforcing equality. Note that if all variables are equal, they are also equal to their average. The PH algorithm incrementally enforces non-anticipativity by penalizing deviations from the average values of decision variables.

PH starts solving each scenario independently, as follows:

$$\min_x f(x, s)$$
$$s.t.x \in C_s$$

PH then computes a solution average and a convergence value to determine if solutions are sufficiently non-anticipative. The convergence value quantifies the deviation of solutions from the "mean" solution. If the convergence value is sufficiently small, PH terminates because the non-anticipativity constraints are (approximately) satisfied. If not, then PH computes a penalization terms, w, for each decision variable, proportional to both the deviation from the average and a penalty factor ρ. Given these penalization terms that tend to force non-anticipativity, the scenario sub-problems are re-solved. The algorithm iterates until the non-anticipativity constraints are "practically" satisfied.

The following pseudo-code describes the PH algorithm:

1. *Solve each scenario without penalizations*

$$k = 0, \ w_k = 0$$

2. *Compute the average for each variable*

$$\bar{x} = \sum_s \Pr(s)x_s$$

3. *If solutions are sufficiently converged, then exit; else $k = k + 1$*

$$g := \|x - \bar{x}\| < \varepsilon$$

4. *Update penalization terms*

$$w = \rho (x - \bar{x}) + w$$

5. *Solve scenarios with penalization terms*

$$\overline{f_s} (x_s) := f_s (x_s) + w x_s + \frac{\rho}{2} \|x_s - \bar{x}\|^2$$

6. *Goto step 2*

If all decision variables are continuous, the PH algorithm provably converges to an optimal solution (Rockafellar and Wets 1991). In the problem we consider, with integer variables, no such convergence proof exists. However, several enhancements to the basic PH algorithm have been introduced to ensure and accelerate convergence in the mixed-integer case (Watson and Woodruff 2010). Specifically, the following issues were carefully addressed in our implementation of PH:

1. **The quadratic term complicates the solution procedure**: In particular, quadratic mixed-integer programs are much more difficult to solve than linear mixed-integer programs. To address this issue, we linearized all quadratic penalty (proximal) terms. In the case of linear decision variables, the number of pieces used in the linearization must be determined. In the case of binary decision variables, the quadratic penalty terms simply reduce to linear penalty terms, because they only take on 0–1 values.
2. **The value given to the parameter ρ** (Steps 4 and 5 in the PH pseudo-code): Computational experience has shown that the value given to ρ has a major impact on the convergence of the PH algorithm, and the quality of the resulting solution.
3. **Mipgap tolerance**: When considering each scenario subproblem, there is no need to solve each problem to optimality, particularly in early PH iterations. Allowing the solves for each subproblem to terminate when the solution obtained is within a small optimality gap allows to solve these problems in a much faster way, without significant loss in the quality of final solution.
4. **Hot starts:** A well-known acceleration technique for PH is to use hot starts when resolving scenario subproblems. Specifically, the solution obtained for sub-problem r in iteration k is used as a starting solution for sub-problem r in iteration k + 1. Feasibility is guaranteed, as only the optimization objective is modified.
5. **Strengthening of the formulation.** Several such strengthenings were implemented, as discussed above in Sect. 4. These strengthenings lead to significantly faster solution times for scenario subproblems.
6. **Variable fixing**. As noted, the PH algorithm is used as a heuristic for our problem due to the presence of 0–1 variables. To obtain solutions more quickly, our implementation fixes individual discrete variables when they have satisfied non-anticipativity for a number of iterations. Fixing variables in such a manner may lead to sub-optimal solutions, but computational experience supports the overall effectiveness of this strategy. Experiments were carried out to determine how many iterations should be considered with common values in the variables before fixing them.
7. **Early termination of PH**. There is no guarantee of convergence for the class of non-convex problems we consider, which include discrete decision variables. Furthermore, convergence of the last few variables often requires more computational effort than it is

worth. Computational experience has shown it is valuable, after a number of PH iterations, to switch the solution algorithm to a regular branch and cut scheme. Specifically, a commercial code is used to solve the full extensive form of the stochastic optimization problem, subject to fixing variables whose values have converged. The solution time for the branch-and-cut approach applied to the extensive form is now considerably faster since a significant number of the 0–1 variables are already fixed.

The combination of these factors has enabled implementation of a PH algorithm that is efficient, with relatively good quality of solutions. The primary potential advantage of PH over direct solution of either the explicit or implicit extensive form lies in scalability, when very large numbers of scenarios are considered. In this case, the extensive form is often computationally intractable for commercial deterministic solvers, in terms of memory, run-time, or both.

6 The test case

The case we are dealing with is the Los Copihues forest, shown in Fig. 1. It has 25 stands, with a total area of 300 hectares, 150,000 m^3 of timber, 15 existing and 11 potential roads, 9 origins where timber is extracted, 3 road intersections, and one exit to a public road. A discount factor of 10 % was used because it is consistent with commonly used values in the Chilean forestry industry to capture the risk free rate and adjustments for risk. The experiments were done under Linux using a computer with Intel(R) Core(TM) 2 Quad CPU Q8200 @ 2.33GHz and four gigabytes of main memory.

To generate the scenarios we first consider market uncertainties. We represent the market uncertainty exactly as in Quinteros et al. (2009), with two random variables: prices and quantity demanded. First, price is simulated as a random walk, and second, the quantity demanded is assumed to have a positive correlation with the price considering the bounds on demand. Then, we also develop scenarios considering uncertainty in timber growth. To generate the timber growth component of the scenarios we considered an average of yearly growth of 3 % with a standard deviation of 0.15 %, which correspond to typical observed values in timber growth in this region of Chile based on conversations with foresters.

In general there are markets for timber products, and Chile has a small share of the world market, so it could be argued that uncertainty should only be expressed in terms of prices. However, for Chilean forest firms, the market is not completely fluid. There are contracts, existing and future ones that may be signed for given quantities. There are markets that need to be supplied with a minimum quantity to be retained, and others where the firm has saturated its ability to supply. In addition, future competitor producers may appear, like Russia increasing significantly its production, which might make some markets unavailable, due to shipping costs. These conditions influenced us to include also uncertainty in future demand. We modeled minimum quantities for retention of a customer and maximum quantities to avoid saturation by putting upper and lower bounds on production. A more complex scheme would be to have multiple prices and various demand breakpoints, but we lacked the data to set these values and this is beyond the scope of the present work.

For our problem, we added uncertainty in growth and yield, that is, the m^3 per hectare obtained when harvesting. Variability is assumed completely correlated between stands, (the case of less correlation does not alter the solution approach), and variation in timber production is assumed to be uniformly distributed over plus and minus 10 % of average

Table 1 Problem sizes for the implicit and explicit extensive forms (EF)

Scenarios	1	18	64	144	162	216	324
	Implicit EF						
Binary cols	156	1, 209	3, 315	7, 410	9, 867	10, 920	16, 185
Linear cols	84	651	1, 785	3, 990	5, 313	5, 880	8, 715
All columns	240	1, 860	5, 100	11, 400	15, 180	16, 800	24, 900
Rows	179	2, 812	9, 746	21, 916	25, 048	32, 824	49, 186
Non-zeros	1, 860	10, 844	38, 052	85, 592	97, 076	128, 288	192, 332
	Explicit EF						
Binary cols	156	2, 808	9, 984	22, 464	25, 272	33, 696	50, 544
Linear cols	84	1, 512	5, 376	12, 096	13, 608	18, 144	27, 216
All columns	240	4, 320	15, 360	34, 560	38, 880	51, 840	77, 760
Rows	179	5, 682	21, 716	48, 936	52, 698	73, 704	110, 856
Non-zeros	1, 860	16, 584	61, 992	139, 632	152, 376	210, 048	315, 672

value. Then, we combine the 16 market scenarios with production variability, creating up to 324 scenarios, which a significantly larger number than previously reported.

For example, a case with 144 scenarios in 4 periods, was created combining market and tree growth scenarios. The scenarios have a tree form, as shown in Fig. 2, with branches emerging from a root node that corresponds to the present period and then from continuing nodes in the tree. One example has (1,3,3,2) branchings for market price, and (1,2,2,2) branching for tree growth. In this case the total number of scenarios becomes $1 \times 3 \times 3 \times 2 \times 1 \times 2 \times 2 \times 2 = 144$ scenarios. In general the number of branches is chosen with a larger number of possibilities in the early periods.

As the number of scenarios grows, so does the dimension of the model. Table 1 shows the dimension for both the explicit and implicit formulations.
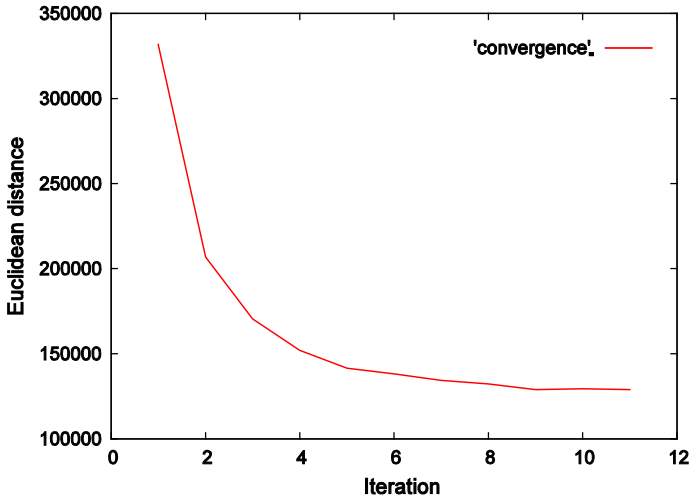
7 Results

Several options were considered in the problems runs.

(a) Use of hot starts
(b) Linearization of quadratic penalization
(c) Strengthening of the formulation
(d) Use of gaps for early termination of MIP problems
(e) Value assigned to parameters ρ and ϵ for termination
(f) Fixing variables

The effect of strengthenings depends on the instance, the solver, other options, and other strengthenings used. When combined with the best set of options for our experiments, the best set of strengthenings we found to be Flow Balance in the Network, Lifting on Road Capacities, Road Flow Capacity.

Hot-starts provided 11 % faster solutions. After testing, a gap of 2 % was found to be the best compromise between computational effort and solution quality. The best results were found in linearizing the quadratic penalization term into just two segments, in V form. For binary variables, the quadratic term is simply approximated linearly. Fixing variables, setting

Fig. 3 Example of a convergence of PH; a 144 scenario example

Table 2 Computational results obtained by varying the number of scenarios

Scenarios	Cplex EF value (1 h)	Cplex EF gap (1 h) (%)	PH+ value	PH+ gap (%)	PH+ run time	PH+ versus Cplex value (%)	Total fixed variables
18	$4,928,180	0.31	$4,920,078	0.47	4 min 22 s	−0.16	2
64	$5,357,780	1.29	$5,386,971	0.74	13 min 1 s	0.54	1719
144	$5,266,830	2.12	$5,287,935	1.68	21 min 42 s	0.40	5051
162	$5,187,040	3.10	$5,242,032	1.98	22 min 54 s	1.05	823
216	$5,332,550	3.92	$5,437,714	1.87	37 min 11 s	1.93	1828
324	$5,545,260	2.91	$5,536,196	2.99	71 min 39 s	−0.16	740

ρ as in Watson et al. (2012), and early termination proved to increase efficiency in solving the problems. We fixed $\epsilon = 0.01$ and $k_{max} = 10$, based on preliminary experiments. Figure 3 shows the convergence graph for 144 scenario case. Convergence is not always monotonic, but the graph gives the reader a sense of typical convergence behavior.

The results for diverse instances are shown in Table 2, where a comparison between solving the extensive form in CPLEX 10.2 with default setting and solving the problem using PH with the same CPLEX as the subproblem solver. We see that for 18 scenarios EF alone obtains slightly better solutions in more time and for this particular set of 324 scenarios where PH+ is slightly dominated by the EF. For all other scenario sets, PH+ dominates the EF alone.

Parallelization of PH shows promise for extremely large numbers of scenarios. Descriptions of parallel implementation and experimentation are beyond the scope of this paper. However, to give the reader a sense of the impact of parallel PH, we conducted an experiment on the 64-scenario case and obtained an objective function value of 5,347,334.56. For this experiment, we parallelized by forming bundles of 4 scenarios each, which we treated as "super-scenarios". If these super-scenarios are solved in parallel, then only a couple of minutes of wall clock time is required for solution. A full study of parallelization on larger instances remains as future work.

8 Conclusions

Solutions for optimization models for medium-term forest management that combine road and harvest planning can be obtained for realistic sized problem instances, even in the face of uncertainty. In this paper we have described strengthening methods that enable solution of problems with many more scenarios previously reported in the literature. Without these strengthenings, solution of the extensive form was limited to very small instances. Furthermore, we have demonstrated that Progressive Hedging is competitive even on a modest serial computer (in our case, a standard laptop). This is important because PH can be parallelized to solve instances with even more scenarios where the attempts to solve the Extensive Form directly will run out of memory. Note that by fixing large numbers of integer variables, the PH algorithm produces a modified extensive form that is smaller and that can be solved with fewer branch and bound nodes. For a particular instance in practice, there is a tradeoff between using more scenarios to get better representations of the uncertainty and the need for more processors and more time for PH to reduce the number of variables so that the resulting modified EF can be solved.

The comparisons we made were done using the best version of CPLEX available at the time. Naturally, CPEX is constantly improving, so the number of scenarios that can be solved using the extensive form is increasing and so is the number of scenarios that can be solved using PH, which employs CPLEX as a sub-problem solver. In the meantime we have also been improving our code. In particular we are starting to test a parallel version, which has the very significant advantage in that the scenario problems can be run in parallel, rather than serially, which is the present case. All indications are that a parallel implementation will be far superior, in particular for a large number of scenarios. As parallel computers become ubiquitous, the ability of PH to immediately use large numbers of processors will be its chief advantage for practical applications in forestry.

An important avenue of ongoing research concerns improving the scenarios for price (Weintraub and Wets 2013). When generated rigorously, there can be a very large number of scenarios needed to capture a multi-stage progression of wood prices. This increases the importance of the work reported here.

References

Alonso-Ayuso, A., Escudero, L., & Ortuno, M. (2003). BFC, a branch-and-fix coordination algorithmic framework for solving some types of stochastic pure and mixed 0–1 programs. *European Journal of Operational Research*, *151*(3), 503–519.

Andalaft, N., Andalaft, P., Guignard, M., Magendzo, A., Wainer, A., & Weintraub, A. (2003). A problem of forest harvesting and road building solved through model strengthening and Lagrangian relaxation. *Operations Research*, *51*(4), 613–628.

Caroe, C. C., & Schultz, R. (1999). Dual decomposition in stochastic integer programming. *Operations Research Letters*, *24*, 37–45.

Constantino, M., Martins, I., & Borges, J. G. (2008). A new mixed-integer programming model for harvest scheduling subject to maximum area restrictions. *Operations Research*, *56*, 542–551.

Escudero, L. F., Garin, A., Merino, M., & Pérez, G. (2009). On BFC-MSMIP: An exact branch-and-fix coordination approach for solving multistage stochastic mixed 0–1 problems. *TOP*, *17*, 96–122.

Frisk, M., Karlsson, J., & Rönnqvist, M. (2006). RoadOpt: A decision support system for road upgrading in forestry. *Scandinavian Journal of Forest Research*, *21*(Suppl. 7), 5–15.

Goycoolea, M., Murray, A. T., Barahona, F., Epstein, R., & Weintraub, A. (2005). Harvest scheduling subject to maximum area restrictions: Exploring exact approaches. *Operations Research*, *53*(3), 490–500.

Goycoolea, M., Murray, A., Vielma, J. P., & Weintraub, A. (2009). Evaluating approaches for solving the area restriction model in harvest scheduling. *Forest Science*, *55*(2), 149–165.

Guglielmo, L., & Sen, S. (2004). A branch-and-price algorithm for multistage stochastic integer programming with application to stochastic batch-sizing problems. *Management Science*, *50*(6), 786–796.

Guignard, M., Ryu, C., & Spielberg, K. (1994). Model tightening for integrated timber harvest and transportation planning.*Proceedings International Symposium Systems Analysis Management Decisions Forestry, Valdivia, Chile*, (pp. 364–369).

Hof, J., & Pickens, J. (1991). Chance-constrained and chance-maximizing mathematical programs in renewable resource management. *Forest Science*, *7*(18), 308–325.

Jones, J. G., Hyde, J. F. C., III, & Meacham, M. L. (1986). *Four analytical approaches for integrating land management and transportation planning on forest lands* (p. 33). Research Paper INT-361. Ogden, UT: U. S. Department of Agriculture, Forest Service, Intermountain Research Station.

Kirby, M. W., Hager, W., & Wong, P. (1986). Simultaneous planning of wildland transportation alternatives. *TIMS Studies in the Management Sciences*, *21*, 371–387.

Martell, D., Gunn, E., & Weintraub, A. (1998). Forest management challenges for operational researchers. *European Journal of Operational Research*, *104*, 1–17.

Martell, D. (2007). Forest fire management: Current practices and new challenges for operational researchers. In A. Weintraub, C. R. Trond Bjorndal, R. Epstein, & J. Miranda (Eds.), *Handbook of Operations Research in Natural Resources*. New York, NY: Springer Science+Business Media.

McNaughton, A. J., & Ryan, D. (2008). Adjacency branches used to optimize forest harvesting subject to area restrictions on clearfell. *Forest Science*, *4*(13), 442–454.

Murray, A. T., & Church, R. L. (1995). Heuristic solution approaches to operational forest planning problems. *OR-Spektrum*, *17*, 193–203.

Quinteros, M., Alonso, A., Escudero, L., Guignard, M., & Weintraub, A. (2009). Forestry management under uncertainty. *Annals of Operations Research*, *190*, 1572–9338.

Richards, W., & Gunn, A. (2000). A model and Tabu search method to optimize stand harvest and road construction schedules. *Forest Science*, *46*, 188–203.

Rockafellar, R. T., & Wets, R. J.-B. (1991). Scenario and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, *16*, 119–147.

Römisch, A., & Schultz, R. (2001). Multistage stochastic integer programs: An Introduction. In M. Grötschel, S. Krumke, & J. Rambau (Eds.), *Online Optimization of Large Scale Systems* (pp. 581–600). Berlin: Springer.

Watson, J. P., & Woodruff, D. L. (2010). Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. *Computational Management Science*, *8*(4), 355–370.

Watson, J. P., Hart, W., & Woodruff, D. L. (2012). PySP: Modeling and solving stochastic programs in python. *Mathematical Programming Computation*, *4*(2), 109–149.

Weintraub, A. & Wets, R.J.-B. Harvesting management: Generating wood-prices scenarios. Technical Report (2013). Santiago: Systemas Complejos en Ingeneria, Universidad de Chile.

Weintraub, A., & Navon, D. (1976). A forest management planning model integrating sylvicultural and transportation activities. *Management Science*, *22*(12), 1299–1309.

Weintraub, A., & Vera, J. (1991). A cutting plane approach for chance constrained linear programs. *Operations Research*, *39*(5), 776–785.

Weintraub, A., Jones, G. J., Magendzo, A., Meacham, M. L., & Kirby, M. W. (1994). A heuristic system to solve mixed integer forest planning models. *Operations Research*, *42*, 1010–1024.

Weintraub, A., Jones, G. J., Magendzo, A., & Malchuk, D. (1995). Heuristic procedures for solving mixed-integer harvest scheduling-transportation planning models. *Canadian Journal Forest Research*, *25*, 1618–1626.

Wets, R. (1975). On the relation between stochastic and deterministic optimization. In A. Bensoussan & J.L. Lions, (Eds.), *Control theory, numerical methods and computer systems modelling*, Lecture notes in economics and mathematical systems, vol. 107 (pp. 350–361). Berlin: Springer